Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT
Discrete Mathematics 277 (2004) 171–191
DISCRETE MATHEMATICS
www.elsevier.com/locate/disc

On a tiling conjecture of Komlós for 3-chromatic graphs☆

Ali Shokoufandeh[a,*], Yi Zhao[b,1]

[a] Department of Computer Science, Drexel University, Philadelphia, PA 19104, USA
[b] Department of Mathematics, Rutgers University, Piscataway, NJ 08854, USA

Received 8 November 2001; received in revised form 29 July 2002; accepted 16 April 2003

Abstract

Given two graphs G and H, an H-*matching* of G (or a *tiling* of G with H) is a subgraph of G consisting of vertex-disjoint copies of H. For an r-chromatic graph H on h vertices, we write $u = u(H)$ for the smallest possible color-class size in any r-coloring of H. The *critical chromatic number* of H is the number $\chi_{cr}(H) = (r-1)h/(h-u)$. A conjecture of Komlós states that for every graph H, there is a constant K such that if G is any n-vertex graph of minimum degree at least $(1-(1/\chi_{cr}(H)))n$, then G contains an H-matching that covers all but at most K vertices of G. In this paper we prove that the conjecture holds for all sufficiently large values of n when H is a 3-chromatic graph.

Keywords: Extremal graph theory; Tiling; Regularity lemma; Blow-up lemma; Critical chromatic number

1. Introduction

All graphs considered in this paper are finite, undirected and simple. If H is a graph on h vertices and G is a graph on n vertices, the objective of *tiling* problems in extremal graph theory is to find many vertex disjoint copies of H in G, or even a complete tiling (called H-factor) of G with $\lfloor n/h \rfloor$ copies of H. One of the earliest tiling results is Dirac's theorem on Hamilton paths [5] that solves the 1-factor problem.

☆ The work is supported in part by National Science Foundation grant DMS-9801396.
* Corresponding author.
E-mail addresses: ashokouf@cs.drexel.edu (Ali Shokoufandeh), zhao@math.uic.edu (Yi Zhao).
1 Present address: Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60607. The research of this author was supported in part by a DIMACS (Center for Discrete Mathematics and Theoretical Computer Science) graduate student Fellowship.

doi:10.1016/S0012-365X(03)00157-2

The case of triangle-factors is due to Corrádi and Hajnal [4], and the celebrated result of Hajnal and Szemerédi settles the K_r-factor problem for all r:

Theorem 1 (Hajnal and Szemerédi [7]). *Let G be a graph on n vertices with minimum degree*

$$\delta(G) \geqslant \left(1 - \frac{1}{r}\right) n,$$

then G has a K_r-factor.

During the 1990s, Alon and Yuster extended the Hajnal–Szemerédi theorem in various ways:

Theorem 2 (Alon and Yuster [2]). *For every $\varepsilon > 0$ and for every integer h, there exists an $n_0 = n_0(\varepsilon, h)$ such that for every graph H on h vertices with chromatic number $\chi(H)$, any graph G with $n > n_0$ vertices and with minimum degree*

$$\delta(G) \geqslant \left(1 - \frac{1}{\chi(H)}\right) n \tag{1}$$

contains at least $(1 - \varepsilon)n/h$ vertex disjoint copies of H.

Theorem 3 (Alon and Yuster [3]). *For every $\varepsilon > 0$ and for every integer h there exists an $n_0 = n_0(\varepsilon, h)$ such that for every graph H on h vertices and for every $n > n_0$, any graph G with n vertices and minimum degree*

$$\delta(G) \geqslant \left(1 - \frac{1}{\chi(H)} + \varepsilon\right) n \tag{2}$$

has an H-factor.

They conjectured that two error terms in above theorems ($\varepsilon n/h$ in Theorem 2 and εn in Theorem 3) could be relaxed to a constant. In [2] they also remarked that this is essentially best possible. These conjectures have been recently proven by Komlós et al. [10]:

Theorem 4 (Komlós et al. [10]). *For every graph H there is a constant K such that if G is an n-graph satisfying*

$$\delta(G) \geqslant \left(1 - \frac{1}{\chi(H)}\right) n, \tag{3}$$

then it has an H-matching that covers all but at most K vertices.

Theorem 5 (Komlós et al. [10]). *Given the conditions of Theorem 4, if*

$$\delta(G) \geqslant \left(1 - \frac{1}{\chi(H)}\right) n + K, \tag{4}$$

then G has an H-factor.

Let us use the notation

$$TT(n,H)=\min\{t:\delta(G)\geqslant t \text{ implies that } n\text{-graph } G \text{ has an } H\text{-factor}\}$$

and define $TT(n,H,M)$ to be the smallest integer t such that if G is an n-graph with minimum degree $\delta(G)\geqslant t$, then there is an H-matching covering at least M vertices in G. Then the sharpness of Theorem 2 and Theorem 3 would suggest that the limit of $TT(n,H)/n$ is $1-1/\chi(H)$; hence, just as in Turán-type Theorems, the relevant quantity for tiling problems would also be the chromatic number $\chi(H)$. While this is true for *some* graphs H, it is false for many others: in [8], Komlós presented a much improved form of Theorem 2, and found that for *any* graph H, the crucial quantity for tiling problems is not the chromatic number $\chi(H)$, but the so-called *critical chromatic number* $\chi_{cr}(H)$. For an r-chromatic graph H on h vertices, we write $u=u(H)$ for the smallest possible color-class size in any r-coloring of H. The critical chromatic number of H is the number $\chi_{cr}(H)=(r-1)h/(h-u)$. It is easy to see that $\chi(H)-1<\chi_{cr}(H)\leqslant\chi(H)$, and $\chi_{cr}(H)=\chi(H)=r$ if and only if every r-coloring of H has equal color-class sizes.

Theorem 6 (Komlós [8, lower bound]). *Let H be a graph with parameters $\chi=\chi(H)$ and $\chi_{cr}=\chi_{cr}(H)$. Then, for all $0<M\leqslant n$,*

$$TT(n,H,M)\geqslant M\left(1-\frac{1}{\chi_{cr}}\right)+(n-M)\left(1-\frac{1}{\chi-1}\right). \tag{5}$$

In particular, $TT(n,H)\geqslant(1-1/\chi_{cr})n$.

He also proved a matching upper bound:

Theorem 7 (Komlós [8, upper bound]). *For every graph H and $\varepsilon>0$ there is a threshold $n_0=n_0(H,\varepsilon)$ such that if $n\geqslant n_0$ and G is a graph with n vertices and minimum degree*

$$\delta(G)\geqslant\left(1-\frac{1}{\chi_{cr}(H)}\right)n, \tag{6}$$

then G contains an H-matching that covers all but at most εn vertices.

He also posed the following conjecture:

Conjecture 8 (Komlós [8]). For every graph H, there is a constant $K=K(H)$ such that if G is an n-graph satisfying (6), then G contains an H-matching that covers all but at most K vertices.

This is best possible for every H (by Theorem 5). Hence,

$$\left(1-\frac{1}{\chi_{cr}(H)}\right)n-K\leqslant TT(n,H,n-K)\leqslant\left(1-\frac{1}{\chi_{cr}(H)}\right)n.$$

In this paper we will show that the conjecture holds for all sufficiently large values of n when H is a 3-chromatic graph.

Theorem 9. *For any 3-chromatic graph H on h vertices with $u = u(H)$, there exists an n_0 such that for all $n \geqslant n_0$, if G is any n vertex graph with*

$$\delta(G) \geqslant \left(1 - \frac{1}{\chi_{cr}(H)}\right) n, \tag{7}$$

then G contains an H-matching that covers all but at most $6h(h-u)/u$ vertices of G.

In the proof we will use the concept of *bottle-graphs*. A bottle-graph of chromatic number r is a complete r-partite graph with color-class size vector $(u, w, w, \dots, w)$, where $u = \alpha w$ for some $\alpha \leqslant 1$. Clearly, the critical chromatic number of this graph is $r - 1 + \alpha$. The vector $(\alpha/(r-1+\alpha), 1/(r-1+\alpha), \dots, 1/(r-1+\alpha))$ is called the *color-vector* of the bottle-graph. The parameters u and w are the *neck* and the *width* of the bottle-graph, respectively. Given an r-chromatic graph H of order h with $u = u(H)$, we say that a graph $\mathscr{B} = \mathscr{B}(H)$ is the bottle-graph of H if $\mathscr{B}$ is the smallest bottle-graph with color-vector $\beta = (s, t, \dots, t)$ which contains an H-factor, where $s = u/h$, and $t = (1-s)/(r-1)$. Note, $\chi_{cr}(\mathscr{B}) = \chi_{cr}(H) = 1/t = (r-1)/(1-s)$. We can always construct a bottle-graph using $r-1$ vertex disjoint copies of H: given color-class sizes $u, u_1, u_2, \dots, u_{r-1}$ in a coloring of H, the ith copy of H places its $u, u_i, u_{i+1}, \dots, u_{r-1}, u_1, \dots, u_{i-1}$ vertices into color classes $1, 2, \dots, r$ of the bottle-graph, respectively. Thus the order of $\mathscr{B}(H)$ is at most $(r-1)h$. Therefore, it is sufficient to prove Theorem 9 for bottle-graphs:

Theorem 10 (Main theorem). *For a graph $H = K(u, w, w)$, there exists an n_0 such that for all $n \geqslant n_0$, if G is any graph on n vertices with minimum degree*

$$\delta(G) \geqslant \frac{\alpha+1}{\alpha+2} n = \left(\frac{1}{2} + \gamma\right) n, \tag{8}$$

where $\alpha = u/w$ and $\gamma = \alpha/2(\alpha+2)$, then G contains an H-matching that covers all but at most $3/\gamma w$ vertices.

2. Notations and tools

$V(G)$ and $E(G)$ denote the vertex-set and the edge-set of the graph G, and we write $v(G) = |V(G)|$ (order of G) and $e(G) = |E(G)|$ (size of G). $N(v, X)$ is the set of neighbors of $v \in V$ in the set $X \subset V$. We use $N(v)$ to denote $N(v, V)$. Hence, $|N(v, X)| = \deg(v, X) = \deg_G(v, X)$ is the degree of v in X, and $\deg(v) = \deg(v, V)$. In a directed graph D, we use $N_{out}(v)$ to denote $\{u \in V(D) |\ (v, u) \in E(D)\}$ (the *out* neighborhood of v), and $\deg_{out}(v) = |N_{out}(v)|$. $\delta(G)$ stands for the minimum degree, and $\Delta(G)$ stands for the maximum degree in G. $v_i(G)$ denotes the size of a maximum set of vertex disjoint i-stars (stars with i leaves) in G. We write $\chi(G)$ and $\chi_{cr}(G)$ for the chromatic number and critical chromatic number of G, respectively. For an r-chromatic graph H of order h, we write $u = u(H)$ for the smallest possible color-class size in any r-coloring of H. When A and B are disjoint subsets of $V(G)$, we use $\deg(A, B)$ to denote the number of edges in $E(G)$ with one endpoint in A and the other in B. A bipartite graph G with color classes A and B and edge set E will be denoted by

$G=(A,B,E)$. $K(n_1,n_2,\ldots,n_r)$ is the complete r-partite graph with color class sizes $n_1,n_2,\ldots,n_r$. The *density* between disjoint sets X and Y is defined as:

$$d(X,Y)=\frac{\deg(X,Y)}{|X||Y|}.$$

In the proof of the Main Theorem, Szemerédi's *Regularity Lemma* [13] plays a pivotal role. We will need the following definition to state the regularity lemma:

Definition 11 (Regularity condition). Assume $\varepsilon>0$. A pair (A,B) of disjoint vertex-sets in G forms an ε-*regular* pair if for every $X\subset A$ and $Y\subset B$ satisfying

$$|X|>\varepsilon|A|,\ |Y|>\varepsilon|B|,$$

we have

$$|d(X,Y)-d(A,B)|<\varepsilon.$$

This definition implies that regular pairs are highly uniform bipartite graphs; namely the density of any reasonably large subgraph is almost the same as the density of a regular pair. We will use the following form of the Regularity Lemma:

Lemma 12 (Degree form). *For every $\varepsilon>0$ there is an $M=M(\varepsilon)$ such that if $G=(V,E)$ is any graph and $d\in[0,1]$ is any real number, then there is a partition of the vertex set V into $\ell+1$ clusters $V_0,V_1,\ldots,V_\ell$, and there is a subgraph G' of G with the following properties*:

- $\ell\leqslant M$,
- $|V_0|\leqslant\varepsilon|V|$,
- *all clusters V_i, $i\geqslant 1$, are of the same size $L\leqslant\lceil\varepsilon|V|\rceil$,*
- $\deg_{G'}(v)>\deg_G(v)-(d+\varepsilon)|V|$ *for all* $v\in V$,
- $G'|_{V_i}=\emptyset$ (V_i *is an independent set in* G'), *for all* i,
- *all pairs (V_i,V_j), $1\leqslant i<j\leqslant\ell$, are ε-regular, each with density either* 0 *or greater than d, in G'.*

A stronger one-sided property of regular pairs is super-regularity:

Definition 13 (Super-regularity condition). Given a graph G and two disjoint subsets of its vertices A and B, the pair (A,B) is (ε,d)-*super-regular* if it is ε-regular and

$$\deg(a)>d|B|\quad\forall a\in A,\qquad \deg(b)>d|A|\quad\forall b\in B.$$

We also use the Blow-up Lemma (see [9,11]):

Lemma 14. *Given a graph R of order r and positive parameters δ and Δ, there exists an $\varepsilon=\varepsilon(\delta,\Delta,r)>0$ such that the following holds. Let $n_1,n_2,\ldots,n_r$ be arbitrary positive integers, and let us replace the vertices $v_1,v_2,\ldots,v_r$ of R with pairwise disjoint sets $V_1,V_2,\ldots,V_r$ of sizes $n_1,n_2,\ldots,n_r$ (blowing up). We construct two graphs on the same* *vertex-set $V = \bigcup_i V_i$. The first graph R_b is obtained by replacing each edge $\{v_i, v_j\}$ of R with the complete bipartite graph between the corresponding vertex-sets V_i and V_j. A sparser G is constructed by replacing each edge $\{v_i, v_j\}$ of R arbitrarily with some (ε, δ)-super-regular pair between V_i and V_j. If a graph H with $\Delta(H) \leqslant \Delta$ is embeddable into R_b, then it is already embeddable into G.*

3. Number of left-over vertices

In Conjecture 8, K is a constant that depends only on H. El-Zahar's conjecture [6] proved recently by Abbasi [1], states that $K(C_l) = 0$. But in general K may not be zero. The following example justifies the need for the leftover vertices, i.e., for some graph H with $\chi_{\rm cr}(H) < \chi(H)$, there exists a constant $K > 0$ such that $TT(n, H, n - K) > (1 - 1/\chi_{\rm cr}(H))n$. It also suggests that the number of leftover vertices in Theorem 10 is correct, up to a constant factor.

Let G be an n-graph satisfying:

(1) $V(G) = A \cup B$, where $|A| = (\frac{1}{2} - \gamma)n$ and $|B| = (\frac{1}{2} + \gamma)n$,
(2) $E(G)$ is complete between A and B,
(3) A is an independent set, B contains $s = \lfloor |B|/(4\gamma n + 2w - 2) \rfloor$ connected components, each of which is a complete bipartite graph (L_i, R_i).
(4) $|L_i| = |R_i| = 2\gamma n + w - 1$ for $1 \leqslant i \leqslant s - 1$.

It is easy to see that G satisfies the degree condition in (8). Our H is again $K(u, w, w)$. Without loss of generality, we assume $w + u$ divides $2\gamma n$.

Let $\mathscr{H}$ denote an H-matching of G. For every copy of H in $\mathscr{H}$, two of its color classes have to come from an (L, R) component of B, and the third color class should be in A. Let us assume that every copy of H in $\mathscr{H}$ has one of its w-vertex classes (*called w-class*) in A, *i.e.*, restriction of $\mathscr{H}$ on B is a union of $K(w, u)$ graphs. Then there are at least $2w - 2$ vertices left uncovered in each component (L_i, R_i) of B, for $i = 1, \ldots, s - 1$. In fact, if we let $t = 2\gamma n/(w + u)$, then we can find $2t$ copies of $K(w, u)$ in each (L, R) component by placing t copies of w- and u-classes in each side (thus $2w - 2$ vertices will be left uncovered). Assume to the contrary that $2t + 1$ copies of $K(w, u)$ could be placed in an (L, R) component of B, say, L contains $t + i$ w-sets and $t - i + 1$ u-sets with $i \geqslant 1$. But this is impossible, because

$$(t + i)w + (t - i + 1)u = tw + tu + i(w - u) + u \geqslant t(w + u) + w > 2\gamma n + w - 1.$$

Finally, our assumption that all copies of H have one of their w-sets in A is necessary for achieving the maximum number of copies of H in $\mathscr{H}$. In fact, embedding $K(w, w)$ in the (L, R) of B will generate less than $2t$ copies of H in (L, R, A). Therefore, in any H-matching of G, there will be at least $2w - 2$ uncovered vertices for each component $(L_i, R_i) \in B$, $1 \leqslant i \leqslant s - 1$, leaving a total of at least $(s - 1)(2w - 2)$ vertices uncovered in B. Consequently, there will be at least $(s - 1)(2w - 2)|A|/|B|$ vertices left uncovered in A. Together, the number of uncovered vertices in any H-matching of G is at least

$$(s-1)(2w-2)\frac{n}{|B|} = \left\lceil \frac{|B|}{4\gamma n + 2w - 2} \right\rceil \frac{n}{|B|}(2w-2) \approx \frac{w}{2\gamma}.$$

4. Proof of the main theorem

4.1. Outline of the proof

In the test graph $H = K(u,w,w)$, we assume $u < w$ (that is, $\chi_{cr} < \chi$), since otherwise an almost-complete tiling follows from Theorem 4, which gives even a better constant. Throughout the paper, we assume that n is sufficiently large and will use the following main parameters:

$$\varepsilon \ll d \ll \sigma \ll \mu \ll 1 - \alpha. \quad (9)$$

For simplicity, we do not compute the actual dependencies among these parameters.

We first apply Lemma 12 to G, with ε and d as in (9), to get a partition of $V(G)$ into clusters $V_0, V_1, \ldots, V_\ell$. Without loss of generality, we assume that L, the size of V_i, $1 \leqslant i \leqslant \ell$, is divisible by $2u(w-u)(w+u)(2w+u)$, because otherwise we could move a constant number of vertices from each V_i to V_0 to achieve this condition. We let $G'' \subset G'$ be the induced graph on $V(G') \setminus V_0$ and define the *reduced graph* R as follows:

The vertices of R are the clusters V_i, $1 \leqslant i \leqslant \ell$, and there exists an edge between two vertices of R if the corresponding clusters form an ε-regular pair in G' with density exceeding d. Since $\delta(G') \geqslant (\frac{1}{2} + \gamma - (d+\varepsilon))n$, and $\varepsilon \ll d$, an easy computation shows that:

$$\delta(R) \geqslant (\tfrac{1}{2} + \gamma - 2d)\ell. \quad (10)$$

Throughout the proof we will use two classes of tripartite graphs: the *unbalanced* triangle-graph $H^*(t)$ is one with three color-classes of size $t, t, \alpha t$ ($t = cL$, for some constant $0 < c \leqslant 1$), and every pair of its color-classes is ε-regular. The *balanced* triangle-graph $H^{**}(t)$ is defined similarly but all its color classes have the same size t.

The outline of the proof is as follows:

We show that, except for a special case, G'' can be tiled by $H^*(L_1)$ and $H^{**}(L_2)$ in a way that a positive percentage of vertices of G'' are in copies of $H^{**}(L_2)$. Here $L_2 = (2+\alpha)L_1 = cL$ for some constant $0 < c \leqslant 1$. Next, the vertices in V_0 will be inserted into appropriate cluster triangles, H^*'s or H^{**}'s, such that after we remove all copies of H containing new vertices, each of the remaining triangles still has an H-tiling leaving out at most $4w$ vertices. The connection among clusters will finally reduce the number of left-over vertices to a constant that depends only on w and u.

In the special case, we will show that in G, there exists an almost-independent vertex set A of size $n/(2+\alpha)$. Plus, $B = V(G) \setminus A$ can be either tiled by $K(u,w)$ almost completely, or partitioned into B_1 and B_2 with $|B_1| = n/(2+\alpha)$, $|B_2| = \alpha n/(2+\alpha)$, and $d(B_1, B_2) \approx 1$. In either case, we conclude that except for a constant number of vertices, G can be tiled by H.

4.2. The maximal clique cover of the reduced graph

In order to tile G'' with H^* and H^{**}, we need the following preparation in pure graph theory.

Given a graph $\mathscr{G}$, a k-clique cover $\Phi = \{\Phi_k, \Phi_{k-1}, \ldots, \Phi_1\}$ is a collection of disjoint cliques, where Φ_i corresponds to a set of cliques of order i for $i \in \{k, k-1, \ldots, 1\}$, and $V(\mathscr{G}) = \bigcup_{i=1}^{k} V(\Phi_i)$. We will say a k-clique cover $\Phi = \{\Phi_k, \Phi_{k-1}, \ldots, \Phi_1\}$ is *maximal* if for any other k-clique cover $\Phi' = \{\Phi'_k, \Phi'_{k-1}, \ldots, \Phi'_1\}$, if for some $1 \leqslant i \leqslant k$, $|\Phi'_i| > |\Phi_i|$, then there is a $i < j \leqslant k$ such that $|\Phi_j| > |\Phi'_j|$.

Consider a maximal 3-clique cover $\Phi = \{\Phi_3, \Phi_2, \Phi_1\}$ in the graph $\mathscr{G}$. Let K and K' be two cliques in Φ of sizes i and j, with $i \leqslant j$. We say K and K' are:

- *well-connected* (or $K \hookrightarrow K'$) if $\deg(v, K') = j - 1$, $\forall v \in K$,
- *over-connected* (or $K \overset{>}{\hookrightarrow} K'$) if $\deg(K, K') \geqslant i(j-1)$ and KK',
- *under-connected* (or $K \overset{<}{\hookrightarrow} K'$) if $\deg(K, K') < i(j-1)$.

The following propositions hold because Φ is maximal:

Proposition 15. (1) $\deg(c, e) \leqslant 1$, $\deg(e, e') \leqslant 2$, $\deg(\{c, c'\}, \tau) \leqslant 4$ *and* $\deg(\{e, e'\}, \tau) \leqslant 8$ *for any* c, c' *in* Φ_1, e, e' *in* Φ_2, *and* τ *in* Φ_3.

(2) *Fix an edge* $e \in \Phi_2$, *and label its end vertices by* $\mathrm{Top}(e)$ *and* $\mathrm{Bot}(e)$. *If another* $e' \in \Phi_2$ *satisfies* $e \hookrightarrow e'$, *then the vertices of* e' *can be labeled as* $\mathrm{Top}(e')$ *and* $\mathrm{Bot}(e')$ *such that* $\{\mathrm{Top}(e), \mathrm{Top}(e')\} \times \{\mathrm{Bot}(e), \mathrm{Bot}(e')\}$ *form a complete bipartite graph. Moreover, if* $\mathscr{E} = \{e' \in \Phi_2 : e \hookrightarrow e'\}$, *then* $\{\mathrm{Top}(e') : e' \in \mathscr{E}\}$ *and* $\{\mathrm{Bot}(e') : e' \in \mathscr{E}\}$ *form two independent sets.*

(3) *If* $\deg(\{c_1, c_2\}, \tau) = 4$ *for* $c_1, c_2 \in \Phi_1$ *and* $\tau \in \Phi_3$, *then* $c_1 \hookrightarrow \tau$, $c_2 \hookrightarrow \tau$, *and* $N(c_1, \tau) = N(c_2, \tau)$. *If* $\mathrm{Top}_c(\tau)$ *denotes the vertex of* τ *that is not contained in* $N(c, \tau)$ *when* $c \hookrightarrow \tau$, *then* $\mathrm{Top}_{c_1}(\tau) = \mathrm{Top}_{c_2}(\tau)$ (*we can simply use* $\mathrm{Top}(\tau)$ *to denote this vertex, as it is independent of the choice of vertex* $c \in \Phi_1$). *It is obvious that* $\mathrm{Top}(\tau)$ *plays the same role as the elements of* Φ_1, *and* τ *can be replaced by* $c_1 \cup N(c_2, \tau)$. *Moreover, if* $\mathscr{S}_1 = \{\tau \in \Phi_3 : c_1 \hookrightarrow \tau, c_2 \hookrightarrow \tau\}$, *then* $\{\mathrm{Top}(\tau) : \tau \in \mathscr{S}_1\}$ *is an independent set.*

(4) *If* $\deg(e_1, \tau) = \deg(e_2, \tau) = 4$ *for* $e_1, e_2 \in \Phi_2$ *and* $\tau \in \Phi_3$, *then* $e_1 \hookrightarrow \tau$, $e_2 \hookrightarrow \tau$, *and the same vertex of* τ *is adjacent to both ends of* e_1 *and* e_2. *We will use* $\mathrm{Tip}(\tau)$ *to denote this vertex and it is independent of the choice of edges* e_1 *and* e_2. *Further, we can label the other two vertices of* τ *as* $\mathrm{Top}(\tau)$ *and* $\mathrm{Bot}(\tau)$, *and the end points of* e_i, $i = 1, 2$ *as* $\mathrm{Top}(e_i)$ *and* $\mathrm{Bot}(e_i)$ *such that* $\{\mathrm{Top}(e_1), \mathrm{Top}(e_2), \mathrm{Top}(\tau)\}$ *and* $\{\mathrm{Bot}(e_1), \mathrm{Bot}(e_2), \mathrm{Bot}(\tau)\}$ *are independent sets. Hence, the pair* $\{\mathrm{Top}(\tau), \mathrm{Bot}(\tau)\}$ *plays the same role as* e_1 *or* e_2. *Finally, let* $\mathscr{S}_2 = \{\tau \in \Phi_3 : e_1 \hookrightarrow \tau, e_2 \hookrightarrow \tau\}$, *then* $\{\mathrm{Top}(\tau) : \tau \in \mathscr{S}_2\}$ *and* $\{\mathrm{Bot}(\tau) : \tau \in \mathscr{S}_2\}$ *form two independent sets in* R.

(5) A $\tau \in \Phi_3$ can be over-connected to at most one element in Φ_1 or Φ_2. Moreover, if τ is over-connected to one element, it will be under-connected to any other element of Φ_1 and Φ_2.

As a result of the above properties, we have:

Lemma 16. *Suppose $\Phi = \{\Phi_1, \Phi_2, \Phi_3\}$ is a maximal 3-clique cover of a graph $\mathscr{G}$, with*

$$|\Phi_1| \geqslant 2 \quad \textit{or} \quad |\Phi_1| = 0, \quad |\Phi_2| \geqslant 2.$$

Then

$$|\Phi_3| \geqslant |\Phi_1| + 2\delta(\mathscr{G}) - v(\mathscr{G}).$$

Proof. First, assume that $|\Phi_1| \geqslant 2$. Since Φ is maximal, any two singletons $c, c' \in \Phi_1$ satisfy

$$\deg(\{c, c'\}, \Phi_1) = 0,$$

$$\deg(\{c, c'\}, \Phi_2) \leqslant 2|\Phi_2|,$$

$$\deg(\{c, c'\}, \Phi_3) \leqslant 4|\Phi_3|.$$

Together we have

$$2\delta(\mathscr{G}) \leqslant \deg(\{c, c'\}, V(\mathscr{G})) \leqslant 2(2|\Phi_3| + |\Phi_2|).$$

Combining with the fact that $3|\Phi_3| + 2|\Phi_2| + |\Phi_1| = v(\mathscr{G})$, the claim thus follows. Next, assume $|\Phi_1| = 0$, $|\Phi_2| \geqslant 2$. Any two edges $e, e' \in \Phi_2$ satisfy

$$\deg(\{e, e'\}, \Phi_2 \setminus \{e, e'\}) \leqslant 4(|\Phi_2| - 2),$$

$$\deg(\{e, e'\}, \Phi_3) \leqslant 8|\Phi_3|,$$

and consequently

$$4\delta(\mathscr{G}) - 8 \leqslant \deg(\{e, e'\}, V(\mathscr{G}) \setminus \{e, e'\}) \leqslant 4(|\Phi_2| - 2) + 8|\Phi_3|.$$

Using $3|\Phi_3| + 2|\Phi_2| = v(\mathscr{G})$ and $|\Phi_1| = 0$, the claim follows. □

We now find a maximal 3-clique cover $\Phi = \{\Phi_3, \Phi_2, \Phi_1\}$ in the reduced graph R and use φ_i to denote the normalized size $|\Phi_i|/\ell$, for $1 \leqslant i \leqslant 3$. Each element (clique) of Φ corresponds to a *cluster-clique* in G''. Corresponding to $\Phi = \{\Phi_3, \Phi_2, \Phi_1\}$, the family of cluster-cliques is denoted by $\Psi = \{\Psi_3, \Psi_2, \Psi_1\}$.

We assume that at least one of $|\Phi_2|$ and $|\Phi_1|$ is bigger than one. Otherwise we remove all the vertices (of G) in Ψ_1 and Ψ_2 to V_0 such that the remaining vertices of G'' are covered by the copies of $H^{**}(L)$ (from Ψ_3) and still $|V_0| \leqslant 4\varepsilon n$. This helps us jump to Section 4.3.2. Furthermore, if $|\Phi_1| = 1$ and $|\Phi_2| > 1$, we remove all the vertices in Ψ_1 to V_0 such that the resulting cover Φ satisfies $|\Phi_1| = 0$. Hence, we can apply Lemma 16 to R and Φ. Using the degree condition (10), we get

$$\varphi_3 \geqslant \varphi_1 + 2\gamma - 4d. \tag{11}$$

Throughout Section 4.3, we will assume that $\varphi_3 > \varphi_1 + 2\gamma + \sigma$ for some positive number σ defined in (9). The special case $\varphi_3 \leqslant \varphi_1 + 2\gamma + \sigma$ will be discussed in Section 4.4.

4.3. The general case

Let $s = \varphi_3 - \varphi_1 - 2\gamma$. Our assumption is that $s > \sigma$.

4.3.1. The decomposition lemma

For two cliques K and K' in the reduced graph R, we say that K' is *good* for K, if $K \hookrightarrow K'$ or $K \overset{>}{\hookrightarrow} K'$, otherwise K' is *bad* for K. An element x of Φ_1 or Φ_2 is called *typical* if for the constant $b = \lceil (1+\alpha)/(1-\alpha) \rceil$, $|\{\tau \in \Phi_3\colon x \overset{>}{\hookrightarrow} \tau\}| < b$, otherwise, x is referred to as an *atypical* element. The same terminology will be used in G'' as well.

First let us estimate the numbers of good triangles in Φ_3 for a given typical edges e in Φ_2. Let $\lambda_e = |\{\tau \in \Phi_3 : \deg(e, \tau) \leqslant 3\}|/\ell$. Then

$$2(\tfrac{1}{2} + \gamma - 2d)\ell \leqslant \deg(e, R) \leqslant (4\varphi_3 - \lambda_e + 2\varphi_2 + \varphi_1)\ell + 2b, \tag{12}$$

which implies $\lambda_e \leqslant \varphi_1 + s + 5d$. Consequently, there are at least $(2\gamma - 5d)\ell$ good triangles in Φ_3 for e.

Similarly, for a typical singleton $c \in \Phi_1$, let $\lambda_c = |\{\tau \in \Phi_3 : c \overset{<}{\hookrightarrow} \tau\}|/\ell$. By the degree condition and the fact that Φ is maximal, we have

$$(\tfrac{1}{2} + \gamma - 2d)\ell \leqslant (2\varphi_3 - \lambda_c + \varphi_2)\ell + b,$$

which implies $\lambda_c \leqslant s/2 + 3d$. Consequently, there are at least $(\varphi_1 + 2\gamma + s/2 - 3d)\ell$ good triangles for each typical $c \in \Phi_1$.

The Slicing Lemma in [12] says that subgraphs of a regular pair are also regular:

Proposition 17. *Let V_i and V_j , $1 \leqslant i, j \leqslant \ell$, be two clusters in G'' that correspond to endpoints of an edge e in the reduced graph R. If both of V_i and V_j are partitioned to p sub-clusters $\{V_i^1, \dots, V_i^p\}$ and $\{V_j^1, \dots, V_j^p\}$ such that the sizes of sub-clusters are at least cL for some $c \in (0,1)$, then the (V_i^r, V_j^t), $1 \leqslant r, t \leqslant p$, are ε'-regular pairs, with $\varepsilon' = \min(\varepsilon/2, \varepsilon/c)$. In particular, $(V_i^t, V_j^t), 1 \leqslant t \leqslant p$, are p disjoint ε'-regular edges.*

Evenly partitioning both clusters V_i, V_j into p parts thus replaces the old edge (V_i, V_j) with p new edges in the cluster graph. This procedure will be referred to as a *p-partition* of a cluster edge. The p-partition of a cluster triangle is defined similarly. For simplicity, we will still use ε as the parameter in the new regular pairs.

Now we are ready to state the decomposition algorithm of G''.

Decomposition Algorithm:

(1) *u-partition all cluster edges in* Ψ_2:
Recall that u is the neck of the bottle-graph H. Let $L' = L/u$ denote the size of the resulting sub-clusters. A new cluster edge corresponds to the same edge of Φ_2 as before.

(2) *Form copies of* $H^*(L')$ *with new typical cluster edges*:
Suppose (V_1^i, V_2^i) is a new cluster edge corresponding to a typical edge $e \in \Phi_2$. We arbitrarily choose a cluster triangle $T \in \Psi_3$ whose corresponding triangle $\tau \in P_3$ is good for e. We use V_1^i, V_2^i and $\alpha L'$ vertices from the cluster corresponding to $\mathrm{Tip}(\tau)$ to form a copy of $H^*(L')$. We repeat this for all new typical edges in Ψ_2. A cluster triangle could be chosen more than once, but not more than $(1-\alpha)/\alpha u$ times.

(3) *Form copies of* $H^*(L')$ *with new atypical cluster edges*:
Consider a cluster edge (V_1^i, V_2^i) corresponding to an atypical edge $e \in \Phi_2$. By definition, there are at least $\lceil (1+\alpha)/(1-\alpha) \rceil$ triangles in Φ_3 that are over-connected to e. Following Proposition 15.5, the corresponding cluster triangles (is this clear?) were not involved in Step 2. Each of these triangles T has at least two clusters adjacent to both V_1^i and V_2^i; we label one of them as $\mathrm{Tip}(T)$. As in Step 2, we use V_1^i, V_2^i and $\alpha L'$ vertices from $\mathrm{Tip}(T)$ to from a copy of $H^*(L')$.

(4) *Partition all cluster triangles and create a new triangle family*:
Let $T \in \Psi_3$ be a cluster triangle with cluster sizes L, L, and $L - s\alpha L'$ ($0 \leqslant s \leqslant (1-\alpha)/\alpha u$). It means that one of its clusters has been used to form s copies of $H^*(L')$ in either Step 2 or 3. Let $L'' = \alpha L'/(1-\alpha) = L/(w-u)$. It is easy to see that T can be divided into s copies of $H^*(L'')$ and $(1-\alpha)/\alpha u - s$ copies of $H^{**}(L'')$. Repeat this to all cluster triangles and denote by Ψ_3' the new family of $H^{**}(L'')$.

(5) *Form copies of* H^* *with typical clusters in* Ψ_1:
Consider a cluster U corresponding to some typical singleton $c \in \Phi_1$. We choose $(1+\alpha)L/(1-\alpha)L''$ triangles from Ψ_3' whose corresponding triangles in Φ_3 are good for c. Each of these triangles $T = \{V_1, V_2, V_3\}$ contains two clusters V_1 and V_2 adjacent to U. Using $(1+\alpha)/2L''$ vertices from each of V_1, V_2, and the entire V_3, we make two copies of $H^*(L''/2)$. The remaining vertices of V_1 and V_2 will be assigned to U. Together with $(1-\alpha)/(1+\alpha)L''$ vertices of U, they form two copies of $H^*((1-\alpha)/2(1+\alpha)L'')$. After repeating this to all selected triangles, we eliminate U.

(6) *Form copies of* H^* *with atypical clusters in* Ψ_1:
By definition, each atypical singleton $c \in \Phi_1$ has at least $\lceil (1+\alpha)/(1-\alpha) \rceil$ over-connected triangles. By Proposition 15.5, these triangles were not involved in Steps 1–5. We thus follow the same procedure in Step 5 to eliminate any cluster corresponding to c.

The correctness of the above algorithm is immediate from the following claim:

Claim 18. *There are enough triangles to carry out steps* 2–6 *of the decomposition algorithm.*

Proof. In Step 1 we created $\varphi_2 u\ell$ sub-cluster edges. To verify the correctness of Step 2, we need to show that when we sequentially consider all new typical edges, there are always good triangle available, even under the constraint that no triangles could be used more than $(1-\alpha)/\alpha u$ times. Recall that the number of good triangles for any typical edge is at least $(2\gamma - 5d)\ell$. We expect the following inequality to hold:

$$\varphi_2 u\ell \leqslant \frac{1-\alpha}{\alpha} u(2\gamma - 5d)\ell.$$

That is, $2\gamma - 5d - \alpha/(1-\alpha)\varphi_2 > 0$. In fact,

$$\begin{aligned} 2\gamma - 5d - \frac{\alpha}{1-\alpha}\varphi_2 &= 2\gamma - 5d - \frac{\alpha(1-3\varphi_3-\varphi_1)}{2(1-\alpha)} \\ &= 2\gamma - 5d - \frac{\alpha}{1-\alpha}\left(\frac{1}{2} - \frac{3}{2}(\varphi_1 + 2\gamma + s) - \frac{1}{2}\varphi_1\right) \\ &= 2\gamma\left(1 + \frac{3\alpha}{2(1-\alpha)}\right) - \frac{\alpha}{2(1-\alpha)} + \frac{2\alpha}{1-\alpha}\varphi_1 \\ &\quad + \frac{3\alpha}{2(1-\alpha)}s - 5d \\ &= \frac{2\alpha}{1-\alpha}\varphi_1 + s_1, \end{aligned} \tag{13}$$

where $s_1 = 3\alpha/2(1-\alpha)s - 5d > \alpha\sigma \gg d > 0$.

In Step 3, u new edges were created from each atypical edge in Φ_2. Observe that the number of (available) over-connected triangles for each atypical edge in Φ_2 is $\lceil (1+\alpha)/(1-\alpha) \rceil$. Again each triangle can be used by $(1-\alpha)/\alpha u$ new edges. The correctness of Step 3 follows from

$$u < \left\lceil \frac{1+\alpha}{1-\alpha} \right\rceil \frac{1-\alpha}{\alpha} u.$$

In Step 4, each triangle in Ψ_3 that has been used by s new cluster edges is partitioned into $L/L'' - s$ copies of $H^{**}(L'')$. After eliminating a total of $L/L'\varphi_2\ell$ new cluster edges, the number of new triangles $|\Psi_3'|$ thus becomes $(\varphi_3 - \varphi_2\alpha/(1-\alpha))L/L''\ell$.

The correctness of Step 6 comes from the same argument as in Step 3. Finally, to finish the proof, we need to justify the correctness of Step 5, or to show that there are enough good triangles in Ψ_3' for all typical singletons of Ψ_1. Because the number of good triangles in Φ_3 is at least $(\varphi_1 + 2\gamma + s/2 - 3d)\ell$, the number of good triangles in Ψ_3' is at least $(\varphi_1 + 2\gamma + s/2 - 3d - \alpha/(1-\alpha)\varphi_2)L/L''\ell$. To guarantee that each single cluster has a disjoint set of $L/((1-\alpha)/(1+\alpha)L'')$ good triangles in Step 5, we need to verify

$$\left(\varphi_1 + 2\gamma + \frac{s}{2} - 3d - \frac{\alpha}{1-\alpha}\varphi_2\right)\ell\frac{L}{L''} - \frac{L}{(1-\alpha)/(1+\alpha)L''}\varphi_1\ell > 0,$$

or equivalently,

$$\varphi_1 + 2\gamma + \frac{s}{2} - 3d - \frac{\alpha}{1-\alpha}\varphi_2 - \left(\frac{1+\alpha}{1-\alpha}\right)\varphi_1 > 0. \tag{14}$$

Substituting for $2\gamma - \alpha/(1-\alpha)\varphi_2 = 2\alpha/(1-\alpha)\varphi_1 + s_1 + 5d$ in (14), we get

$$\frac{s}{2} + s_1 + 2d > \frac{\sigma}{2} > 0. \qquad \square \tag{15}$$

After the decomposition G'' is covered by disjoint copies of H^* and H^{**}, and by (15), at least $(\sigma/2)n$ of its vertices are covered by the copies of H^{**}. It is worth mentioning that in this decomposition there are four possible sizes for copies of H^*: $H^*(L/u)$, $H^*(L/(w-u))$, $H^*(L/2(w-u))$ and $H^*(L/2(w+u))$, while all the copies of H^{**} are $H^{**}(L/(w-u))$. Using Proposition 17, we further partition some cluster triangles such that the resulting cover is made of $H^*(L_1)$ and $H^{**}((2+\alpha)L_1)$, with $L_1 = L/2u(w-u)(w+u)(2w+u)$. Hence,

Lemma 19. *If $\varphi_3 - \varphi_1 - 2\gamma > \sigma$, then G'' can be covered by vertex disjoint copies of $H^*(L_1)$ and $H^{**}(L_2)$, in which $L_2 = (2+\alpha)L_1 = CL$ for some $0 < C \leqslant 1$. Moreover, at least $\sigma/2n$ of the vertices of G are included in the copies of H^{**}.*

4.3.2. Handling of exceptional vertices

The proof of the Main Theorem in the general case is immediate from the following lemma:

Lemma 20. *If a graph G satisfies (8) and contains a vertex subset V_0 of size at most θn, with $\varepsilon \ll \theta \leqslant 1$, and $G'' = G \setminus V_0$ can be partitioned to two disjoint subgraphs $G_1 \cup G_2$ such that:*

(1) *G_1 has an $H^*(L_1)$-factor and G_2 has an $H^{**}(L_2)$, where $L_2 = (2+\alpha)L_1 = CL$, for some constant $0 < C \leqslant 1$,*
(2) *$|V(G_2)| = \rho n$, with $\theta \ll \rho \ll 1$,*

then G has an H-matching leaving at most $3w/\gamma$ uncovered.

Proof. Define $\mathscr{H}^*$ and $\mathscr{H}^{**}$ as the families of H^* and H^{**} used in the tiling of G_1 and G_2, respectively. Let $\mathscr{H} = \mathscr{H}^* \cup \mathscr{H}^{**}$, $h_1 = |\mathscr{H}^*|$, $h_2 = |\mathscr{H}^{**}|$. We know,

$$(2+\alpha)L_1 h_1 + 3(2+\alpha)L_1 h_2 = |V(G'')| = n - \theta n. \tag{16}$$

We assume that $\rho \leqslant \frac{1}{2}$, because we can always reduce the number of H^{**} by dividing one copy of $H^{**}((2+\alpha)L_1)$ into three copies of $H^*(L_1)$. In $T = \{U_1, U_2, U_3\} \in \mathscr{H}^*$, the cluster that contains αL_1 vertices is always denoted by U_3, and referred to as the *small* cluster of T.

We will use two complete tripartite graphs $H_2 = K(w+u, w+u, 2w)$ and $H_3 = K(h, h, h)$. Clearly, both H_2 and H_3 contain H-factors. We also use $H^- = K(w, w, u-1)$ and $H_3^- = K(h, h, h-1)$. The following is a corollary of the Key Lemma in [12]:

Lemma 21. *Suppose $T = \{V_1, V_2, V_3\}$ is an original cluster triangle in G'', i.e., $|V_i| = L$ and (V_i, V_j) is an ε-regular pair. $T' = \{W_1, W_2, W_3\}$ is derived from T with $W_i \subset$*

V_i, $|W_i| = d_1 L$, $i = 1,2,3$, for some $d_1 \geqslant \varepsilon$. Then $H_3 \subset T'$. In particular, $H \subset T'$, its small color class could come from any of W_1, W_2, or W_3.

For a cluster U, we write $v \sim U$ if $\deg(v, U) \geqslant d\,|U|$. For each exceptional vertex $v \in V_0$, we will use vertices in $V(G'')$ to construct either one or three copies of H. There are two possible ways to accomplish this:

- If there exists a triangle $T = \{U_1, U_2, U_3\} \in \mathscr{H}^{**}$ and $i \in \{1,2,3\}$ such that $v \sim U_j$ $\forall j \neq i$, then U_i is called a *host* cluster of v. Let $W_i = U_i$ and $W_j = N(v, U_j)$ for $j \neq i$. Using Lemma 21, we can find a copy of H_3, or an H_3^- in W_1, W_2, W_3. Now $H_3^- \cup \{v\}$ forms a copy of H_3. We remove this H_3 copy such that the resulting U_i has one more vertex than the other two clusters.
- If there exists a triangle $T = \{U_1, U_2, U_3\} \in \mathscr{H}^*$ (U_3 as the small cluster) such that $v \sim U_1$ and $v \sim U_2$, then as above, we can remove an H^- from T (with $u - 1$ vertices from U_3) such that $H^- \cup \{v\}$ forms a copy of H. As a result, U_3 now has one more vertex than $\alpha|U_1|$ and $\alpha|U_2|$.

We sequentially consider all vertices in V_0. For each one, if applicable, we perform one of the above two procedures. Note in above procedures, U_1, U_2, and U_3 always represents *updated* clusters, *i.e.*, they only contains the remaining vertices after some copies of H_3^- or H^- are removed. To prevent a cluster from losing too many vertices, we will leave a cluster-triangle (either in $\mathscr{H}^{**}$ or $\mathscr{H}^*$) alone if any of its clusters has lost $\sqrt{\theta}L_1$ vertices.

When we proceed to the mth vertex of V_0, we need to show that there always exists an available cluster triangle T such that either $T \in \mathscr{H}^*$ satisfies $v \sim U_1$ and $v \sim U_2$ or $T \in \mathscr{H}^{**}$ satisfies $\exists i \in \{1,2,3\}$, $\forall j \neq i, v \sim U_j$.

Let m_1 denote the number of H^* whose small clusters can host v, and m_2 denote the number of H^{**} that contain at least one host cluster for v. Note G'' could have lost at most $(m-1)(3h-1)$ vertices. By the degree condition we have:

$$\begin{aligned}(\tfrac{1}{2} + \gamma - 3h\theta)n &\leqslant (\tfrac{1}{2} + \gamma - \theta)n - (m-1)(3h-1)\\ &\leqslant \deg(v, V(G''))\\ &\leqslant (h_1 - m_1)(1 + \alpha + d)L_1 + m_1(2+\alpha)L_1\\ &\quad + (h_2 - m_2)(1+2d)(2+\alpha)L_1 + 3m_2(2+\alpha)L_1.\end{aligned}$$

We divide last two of these inequalities by $n - \theta n$ and let $t_1 = (m_1(2+\alpha)L_1)/(n-\theta n)$, $t_2 = (m_2 3(2+\alpha)L_1)/(n-\theta n)$. Using $(1+\alpha)/(2+\alpha) = \frac{1}{2} + \gamma$, $\rho = (3(2+\alpha)L_1 h_2)/(n - \theta n)$, and (16), we have

$$\begin{aligned}\frac{1}{2} + \gamma - 3h\theta &\leqslant (1 - \rho - t_1)\left(\frac{1+\alpha+d}{2+\alpha}\right) + t_1 + \left(\frac{1+2d}{3}\right)(\rho - t_2) + t_2\\ &\leqslant \left(\frac{1}{2} + \gamma + d\right)(1-\rho) + \left(\frac{1}{2} - \gamma\right)t_1 + \left(\frac{1}{3} + d\right)\rho + \frac{2}{3}t_2,\end{aligned}$$

and

$$\left(\frac{1}{2}+\gamma-\frac{1}{3}\right)\rho-d-3h\theta \leqslant \frac{1}{2}t_1+\frac{2}{3}t_2 \leqslant \frac{7}{6}\max(t_1,t_2).$$

That implies $\max(t_1,t_2) \geqslant (\frac{1}{7}+6/7\gamma)\rho-d-3h\theta > \frac{1}{7}\rho$. Since $t_1 = m_1(1-\rho)/h_1$, and $t_2 = m_2\rho/h_2$, then either $m_1 > \frac{1}{7}\rho h_1$ or $m_2 > \frac{1}{7}h_1$. Moreover,

$$\min\left(\frac{1}{7}\rho h_1, \frac{1}{7}h_2\right) > \sqrt{\theta}\frac{n}{L_1} = \frac{\theta n}{\sqrt{\theta}L_1}.$$

This means that the above procedure can be repeated for all $v \in V_0$.

After V_0 becomes empty, we add some more vertices to V_0 to achieve super-regularity inside the triangles and then eliminate new (at most $6\varepsilon n$) exceptional vertices. In $T=\{U_1,U_2,U_3\}$, we move a vertex v from cluster U_i to V_0 if there exists $j \neq i$ with $\deg(v,U_j) < (d-\varepsilon)|U_j|$. The ε-regularity of the cluster-pairs guarantees that there are at most $2\varepsilon|U_i|$ such vertices in each cluster. For $T=\{U_1,U_2,U_3\} \in \mathscr{H}^*$, we may also remove extra vertices to maintain $|U_1|=|U_2|$, $|U_3| \geqslant \alpha|U_1|$. After this, all pairs inside the triangles satisfy $(\varepsilon, d/2)$ super-regularity.

The new elements of $\mathscr{H}^*$ do not necessarily follow the original definition of H^*. Instead, if $T=\{U_1,U_2,U_3\}$, we have $|U_1|=|U_2|$ and $|U_3| \geqslant \alpha|U_1|$. For convenience, we still denote such triangles by H^*. Moreover, there might be a discrepancy among three clusters in a triangle H^{**} in $\mathscr{H}^{**}$. Suppose each U_i hosted x_i exceptional vertices, and we set $x=x_1+x_2+x_3$. The current size of U_i is $(2+\alpha)L_1 - xh + x_i$, $1 \leqslant i \leqslant 3$. We divide this triangle into three copies of H^* with clusters sizes $(L(x), L(x), \alpha L(x)+x_3)$, $(L(x), \alpha L(x)+x_2, L(x))$, and $(\alpha L(x)+x_1, L(x), L(x))$, where $L(x)=L_1-wx$. After we repeat this to all the triangles in H^{**}, we obtain a new triangle family (still denoted by $\mathscr{H}$) containing only H^*-type cluster-triangles. Non-small clusters (U_1 or U_2) in different triangles may have different sizes, but they satisfy $|U_1|=|U_2| > (1-\sqrt{\theta})L_1$.

Denote the induced subgraph of G on the remaining vertices by G''. Our goal is to tile G'' with H. Let us first estimate the size of the largest H-matching of a current cluster-triangle.

Claim 22. *For any cluster-triangle $T=\{U_1,U_2,U_3\} \in \mathscr{H}$, there is an H-matching of T that leaves out $K(U_i)$ vertices in U_i for $i=1,2,3$, in which $K(U_1)+K(U_2)+K(U_3) < 4w$.*

Proof. Let $T=\{U_1,U_2,U_3\}$ be a cluster-triangle in $\mathscr{H}$, with $|U_1|=|U_2|=s$, $|U_3|=\alpha s+t$. We will remove i copies of H_2 from T such that in each copy the color class of size $2w$ always resides in U_3. Let U_1', U_2', U_3' denote the new clusters of T. To achieve $|U_3'|=\alpha|U_1'|(=\alpha|U_2'|)$, we want to have

$$\alpha s+t-2iw=\alpha(s-i(w+u)),$$

or

$$i=\frac{t}{2w-\alpha(w+u)}.$$

This, in turn, implies that the above procedure can be repeated until $|U_3'|-\alpha|U_1'| < 2w-\alpha(w+u)$. Finally, by the Blow-up Lemma, the new graph $T'=\{U_1',U_2',U_3'\}$ with

$|U_1'|=|U_2'|=s'$, $|U_3'|=\alpha s'$ can be tiled with H except for at most $h-1$ vertices. Overall there exists an H-matching of T that leaves out at most $h+2w-\alpha(w+u)<4w$ vertices. □

If we directly apply the algorithm in Claim 22 to all the elements of $\mathscr{H}$, we will get an H-matching of G'' that leaves out at most $4w|\mathscr{H}|$ vertices. Since $|\mathscr{H}|=O(\ell)$ and ℓ is not larger than the constant $M(\varepsilon)$ (according to the Regularity Lemma), this already confirms Conjecture 8.

However, using the connection between clusters in different triangles of $|\mathscr{H}|$, the number of uncovered vertices will be reduced to a constant independent of $|\mathscr{H}|$. Instead of performing the tiling immediately, we first use Claim 22 to find the numbers of extra vertices $K(U)$ for each cluster U. We then define a directed graph $\mathscr{D}$ whose vertices are all the clusters (large or small). Suppose $U\in T, U'\in T'$ and $T'=\{U',U'',U'''\}$. We draw a directed edge from U to U' if and only if U is adjacent to both U'' and U'''.

Claim 23. *For any* $C\in V(\mathscr{D})$, $\deg_{\text{out}}(C)\geqslant(4\gamma/3)|V(\mathscr{D})|$.

Proof. For $T=\{U_1,U_2,U_3\}$ in $\mathscr{H}$, if there is a directed edge from C to some cluster in T, C must be adjacent to at least two clusters in T. We define t_1, t_2, and t_3 as the fractions of the triangles $T\in\mathscr{H}$ for which, respectively, C is adjacent to all but U_3, C is adjacent to all but one of U_1,U_2, and C is adjacent to all U_i. By the degree condition, we have

$$\begin{aligned}\tfrac{1}{2}+\gamma-2d &\leqslant (1-2\gamma)t_1+(\tfrac{1}{2}+\gamma)t_2+t_3\\ &=(12-\gamma)+(\tfrac{1}{2}-\gamma)t_1+2\gamma t_2+(\tfrac{1}{2}+\gamma)t_3.\end{aligned}$$

Since $\gamma=\alpha/2(2+\alpha)\leqslant\frac{1}{6}$, we have $(\frac{1}{2}+\gamma)\leqslant 2(\frac{1}{2}-\gamma)$. This in turn implies

$$\begin{aligned}2\gamma-2d &\leqslant (\tfrac{1}{2}-\gamma)t_1+2\gamma t_2+(\tfrac{1}{2}+\gamma)t_3\\ &\leqslant (\tfrac{1}{2}-\gamma)(t_1+t_2+2t_3).\end{aligned}$$

Therefore in $\mathscr{D}$,

$$\begin{aligned}\deg_{\text{out}}(C) &= (t_1+t_2+3t_3)|\mathscr{H}|\\ &\geqslant \frac{2\gamma-2d}{\frac{1}{2}-\gamma}\,\frac{|V(\mathscr{D})|}{3}\\ &> \frac{4\gamma}{3}|V(\mathscr{D})|.\end{aligned}$$ □

In a directed graph D we define the *source set* of a vertex v as

$$\mathscr{W}(v)=\{u\in V(D):\ \text{there is a directed path from } u \text{ to } v\}.$$

It is easy to see that $|\mathscr{W}(u)|\geqslant|\mathscr{W}(v)|$ if $u\in N_{\text{out}}(v)$. We call a set $\mathscr{S}(\mathscr{D})\subseteq V(\mathscr{D})$ the *sink* set of $\mathscr{D}$ if $V(\mathscr{D})=\bigcup_{v\in\mathscr{S}(\mathscr{D})}\mathscr{W}(v)$.

Lemma 24. *In a directed graph $\mathscr{D}$, with $\delta = \min_{v\in V(\mathscr{D})} \deg_{\mathrm{out}}(v)$, there is a sink set $\mathscr{S}$ of size at most $|V(\mathscr{D})|/(\delta+1)$.*

Proof. Let x_1 be a vertex in which

$$|\mathscr{W}(x_1)| = \max_{v\in V(\mathscr{D})} |\mathscr{W}(v)|.$$

Since $N_{\mathrm{out}}(x_1) \in \mathscr{W}(x_1)$ (otherwise, some out-neighbor of x_1 would produce a larger source set), we have $|\mathscr{W}(x_1)| \geqslant \delta+1$. Let $\mathscr{D}' = \mathscr{D} \setminus \mathscr{W}(x_1)$. Then for every vertex $v \in \mathscr{D}'$, the set $N_{\mathrm{out}}(v, \mathscr{D}')$ is the same as $N_{\mathrm{out}}(v, \mathscr{D})$, since if $v' \in N_{\mathrm{out}}(v,\mathscr{D}) \cap \mathscr{W}(x_1)$, v would also be in $\mathscr{W}(x_1)$. Now, let x_2 be a vertex with the largest source set in $\mathscr{D}'$. We have $\mathscr{W}(x_2) \geqslant \delta+1$ in $\mathscr{D}'$. This procedure can be repeated at most $|V(\mathscr{D})|/(\delta+1)$ times and the proof follows. □

We are now ready to describe our tiling algorithm. For each cluster U, by Claim 22, we first find $K(U)$, the number of left-over vertices in each cluster U. Next, applying Claim 23 and Lemma 24 in the directed (cluster) graph $\mathscr{D}$, we find a sink set $\mathscr{S}$ of size at most $3/4\gamma$. We then assume that only the clusters in $\mathscr{S}$ may carry left-over vertices. In fact, assume there exists some cluster $C^0 \notin \mathscr{S}$ with $K(C^0) > 0$. We can find a directed path $C^0, C^1, \ldots, C^t$ from C^0 to some cluster $C^t \in \mathscr{S}$. Set $x = K(C^0)$. For $i = 1, \ldots, t$, suppose T_i is the triangle that contains C^i. Depending on whether C_i is a large or a small cluster, it takes either w or v vertices of C^i to form a copy of H inside T_i. Denote this number by u_i. For $i = 1, \ldots, t$, we form x copies of H with x vertices from C^{i-1} and $(h-1)x$ vertices from T_i (in particular $(u_i - 1)x$ vertices from C^i). After this, $K(C^0)$ becomes zero and $K(C^t)$ is increased by x.

Although one cluster might be included in many such paths, since the total number of extra numbers is much smaller than dL_1, the super-regularity will not be impacted even after the above procedure is applied to all the clusters. Finally, we apply the Blow-up Lemma to all the triangles in $\mathscr{H}$. The only triangles that could carry uncovered cvertices are the ones containing clusters of $\mathscr{S}$, and each of them could carry at most $4w$ uncovered vertices. Therefore, the total number of left-over vertices is at most

$$4w \times \frac{|V(\mathscr{D})|}{4\gamma|V(\mathscr{D})|/3} = \frac{3w}{\gamma}. \qquad \square$$

4.4. The special case

Recall that $\Phi = \{\Phi_1, \Phi_2, \Phi_3\}$ is the maximal clique-cover of R. In this section, we assume that

$$s = \varphi_3 - \varphi_1 - 2\gamma \leqslant \sigma. \tag{17}$$

Depending on the (relative) size of Φ_1, our special case will further be separated into two cases:

4.4.1. Special case (I): $\varphi_1 < \mu$

Since $\sigma \ll \mu \ll 1-\alpha$, we have $\varphi_3 < 2\gamma + 2\mu$ and

$$\varphi_2 = \frac{1}{2}(1-3\varphi_3-\varphi_1) > \frac{1}{2}(1-6\gamma-7\mu) = \frac{1}{2}\left(\frac{2(1-\alpha)}{2+\alpha} - 7\mu\right) \gg \mu. \tag{18}$$

Recall that for any two edges $e, e' \in \Phi_2$, $\deg(\{e,e'\}, \Phi_2) \leqslant 4\varphi_2\ell$, and $\deg(\{e,e'\}, \Phi_3) \leqslant 8\varphi_3\ell$. Let

$$m = |\{\tau \in \Phi_3 : \deg(\{e,e'\}, \tau) < 8\}|/\ell$$

and

$$m' = |\{e \in \Phi_2 : \deg(\{e,e'\}, e) < 4\}|/\ell.$$

Then

$$4(\tfrac{1}{2} + \gamma - 2d) \leqslant 8\varphi_3 - m + 4\varphi_2 - m' + 2\varphi_1,$$

which implies

$$m + m' \leqslant 4\mu + 8d. \tag{19}$$

According to Proposition 15.15, a triangle that is over-connected to some edge must be under-connected to any other edge. According to (19), the number of such triangles is small, and so is the number of edges that are over-connected to some triangles. Since $|\Phi_2|$ is not small, we can always find two edges $e_1, e_2 \in \Phi_2$ such that $\deg(e_i, \tau) \leqslant 4$, for any $\tau \in \Phi_3$ and $i = 1, 2$. Let

$$\mathscr{S} = \{\tau \in \Phi_3 : \deg(\{e_1, e_2\}, \tau) = 8\} = \{\tau \in \Phi_3 : e_1 \hookrightarrow \tau, e_2 \hookrightarrow \tau\}$$

and

$$\mathscr{E} = \{e \in \Phi_2 : \deg(\{e_1, e_2\}, e) = 4\} = \{e \in \Phi_2 : e_1 \hookrightarrow e, e_2 \hookrightarrow e\}.$$

The same computation as above shows that all but at most $4\mu + 8d$ elements of Φ_2 and Φ_3 belong to $\mathscr{S}$ and $\mathscr{E}$. By Proposition 15.2 and 15.4, the sets $A = \{\mathrm{Top}(e), \mathrm{Top}(\tau) : e \in \mathscr{E}, \tau \in \mathscr{S}\}$ and $B = \{\mathrm{Bot}(e), \mathrm{Bot}(\tau) : e \in \mathscr{E}, \tau \in \mathscr{S}\}$ are both independent, and

$$|A| = |B| = \varphi_2 + \varphi_3 - (4\mu + 8d) > \tfrac{1}{2} - \gamma - 6\mu.$$

Let $\beta = 12\mu$. By the degree condition in the reduced graph R for any cluster $a \in A$ and $b \in B$ we have $\deg(a, R \setminus A) > (1-\beta)|R \setminus A|$ and $\deg(b, R \setminus B) > (1-\beta)|R \setminus B|$. Next, we form a set C with clusters $\mathrm{Tip}(\tau)$ for all $\tau \in \mathscr{S}$ and the remaining clusters in R.

From here on, A, B, and C refer to their underlying vertex sets in G. First we move the vertices in V_0 to C. We then remove a vertex $v \in C$ to A (or to B) if $\deg(v, A) < \beta_1|A|$ (or $\deg(v, B) < \beta_1|B|$), where $\beta \ll \beta_1 \ll 1$. We still denote the resulting sets by A, B and C.

In an *ideal* case, $|A| = |B| = 1/(2+\alpha)n$ and $|C| = \alpha/(2+\alpha)n$. By super-regularity between every two class in $\{A, B, C\}$, the Blow-up Lemma produces the desired H-factor in G. If $|A| < 1/(2+\alpha)n$ and $|B| < 1/(2+\alpha)n$, an argument similar to Claim 22 shows that all but $5w$ vertices of G can be covered by disjoint copies of H. Consequently, we may assume that $|A| > 1/(2+\alpha)n$. In the following we consider the case that $|B| < 1/(2+\alpha)n$ and $|C| < 2\gamma n$ (the other cases are similar).

From A we will move a vertex v to one of the classes B or C with fewer vertices for which $\deg(v, A) > \beta_1|A|$. We still denote the resulting sets by A, B and C. After this step, either we can achieve the ideal case, or we have $\deg(v, A) \leqslant \beta_1|A|$, for all $v \in A$. Assume the latter is true and set $t = |A| - 1/(2+\alpha)n$. We need the following fact:

Proposition 25.

$$v_i(G) \geqslant (\delta(G) - i + 1)\frac{n}{2(i+1)\Delta(G)}.$$

To see this, take a maximal set of i-stars in G and let m denote its size. Let $\mathbf{E}$ represent the number of edges between the stars and the remaining vertices of G. We have the following chain of inequalities which proves the proposition:

$$(n - m(i+1))(\delta(G) - (i-1)) \leqslant \mathbf{E} \leqslant m(i+1)\Delta(G).$$

Proposition 25 implies that we can find t vertex disjoint w-stars in A. After moving the centers of these w-stars to either B or C to reach the ideal case, we immediately remove t copies of H that contain these w-stars. The remaining sets A, B and C have size ratio $(1, 1, \alpha)$ and satisfy the super-regularity condition. The Blow-up Lemma completes the proof.

4.4.2. Special case (II): $\varphi_1 \geqslant \mu$

Similarly to Special Case (I), we can find two clusters c, c' in Φ_1 such that no triangle is over-connected to either of c or c'. Let

$$m = |\{\tau \in \Phi_3\colon \deg(\{c, c'\}, \tau) \leqslant 3\}|/\ell \quad \text{and} \quad m' = |\{e \in \Phi_2\colon \deg(\{c, c'\}, e) \leqslant 1\}|/\ell.$$

We have

$$2(\tfrac{1}{2} + \gamma - 2d) \leqslant 4\varphi_3 - m + 2\varphi_2 - m',$$

which implies

$$m + m' \leqslant s + 4d < 2\sigma.$$

Therefore, there are altogether $\varphi_3 + \varphi_2 - 2\sigma$ triangles and edges well-connected to both c and c'. Applying Proposition 15.3 to these triangles τ and edges e, the clusters $\mathrm{Top}(\tau)$ and $\mathrm{Top}(e)$ play the same roles as clusters c and c'. Together with the rest of the clusters in Φ_1, they form an independent cluster set A with

$$|A| \geqslant (\varphi_3 + \varphi_2 + \varphi_1 - 2\sigma)\ell = \left(\frac{1}{2} - \gamma - \frac{\sigma}{2} - 2\sigma\right)\ell.$$

We then define B as the remaining clusters of these triangles and edges, i.e., B is made up of edges and singletons. We add vertices in the clusters which are not in A or B to V_0. Clearly, $|V_0| < 6\sigma n$.

We use V_b for the underlying vertices of G in the clusters of B. Our objective is to cover $V_b \cup V_0$ with copies of $K(w, u)$, then combine each copy of $K(w, u)$ with w vertices of A to obtain the desired H-matching. For simplicity we define H_0 as $K(w, u)$.

To tile $G(V_b \cup V_0)$ with H_0, we almost repeat Section 4.3. The only extra requirement is that vertices in each copy of H_0 must share many common neighbors in A.

Indeed, since B is covered by cluster edges and singletons, as the way we modified $\Psi_3 \cup \Psi_2$ to $\mathscr{H}^{**} \cup \mathscr{H}^{*}$ in Section 4.3.1, we can then cover B with balanced and unbalanced cluster-edges, with cluster ratios (1:1) and $(1:\alpha)$, respectively. This is always possible, because

$$2\gamma - \frac{\alpha}{1-\alpha}\varphi_2 > \frac{2\alpha}{1-\alpha}\varphi_1 = \frac{2\alpha}{1-\alpha}\mu.$$

From here on, A refers to its underlying vertex sets in G. Let $\beta = 6\sigma$. We have

$$(\tfrac{1}{2} - \gamma - \beta)n \leqslant |A| \leqslant (\tfrac{1}{2} - \gamma)n,$$

and

$$\deg(v, V_b) \geqslant (\tfrac{1}{2} + \gamma - 3d)n \geqslant (1-\beta)|V_b|, \quad \forall v \in A.$$

We move at most $2\beta|V_b|$ vertices from V_b to A if such a vertex v satisfies

$$\deg(v, A) < \beta_1 |A|, \tag{20}$$

with $\beta \ll \beta_1 \ll \mu$. To finish the proof of the Main Theorem we need to consider the following three cases:

Case (1): $|A| = (\frac{1}{2} - \gamma)n$

Let $\mathscr{V}$ be a subset of V_b, in which for all $v \in \mathscr{V}$,

$$\deg(v, A) \leqslant (1-\beta_2)|A|, \tag{21}$$

with $\beta \ll \beta_2 \ll \beta_1$.

Let us first consider the ideal case of $|\mathscr{V}| = 0$. We will find an H_0-matching that covers all but a constant number of vertices of $V_b \cup V_0$. Since (21) does not hold, for every copy of H_0 with vertices $\{v_1, \ldots, v_{w+u}\}$, the common neighborhood of v_i, $i = 1, \ldots, w+u$, will almost cover the whole set A. Moreover, the common neighborhood of any w vertices in A is almost $|B|$. Next we break down all but a constant number of vertices of A into sets of size w. By the König–Hall Theorem, there exists a perfect matching between the copies of H_0 from B and the w-sets from A. This in turn implies that G contains disjoint copies of H except for a constant number of vertices.

If no vertices in any clusters of B satisfy (20), we may tile $V_b \cup V_0$ with H_0 exactly as in Section 4.3.2. Otherwise, we have to make sure that the size ratio of two new clusters in any cluster edge fits the need of H_0-tiling. If an unbalanced edge (U_1, U_2) with $|U_1| = L_1 = |U_2|a$ loses some vertices, we remove more vertices from U_1, U_2 to V_0 such that the resulting clusters U_1', U_2' satisfy (I) $|U_1'| > L_1/2$, (II) $1 \geqslant |U_1'|/|U_2'| \geqslant \alpha$. If a balanced edge (U_1, U_2) with $|U_1| = |U_2| = L_2$ loses vertices, we may remove more vertices to V_0 such that the resulting clusters U_1', U_2' satisfy $|U_1'| = |U_2'| \geqslant L_2/2$. It is easy to see that after these steps, the size of V_0 is smaller than $C_0\beta n$, where $C_0 = 4/\alpha$. Next, we follow the same argument as in Lemma 20 to find a H_0-matching of V_b that leaves out only $(1/2+\gamma)/\gamma 2w$ vertices. After combining with A, the total number of vertices uncovered by copies of H is bounded by

$$2w \times \frac{1/2+\gamma}{\gamma} \times \frac{1}{1/2+\gamma} = \frac{2w}{\gamma}.$$

When $|\mathscr{V}| > 0$, we will need an H_0-matching of $V_b \cup V_0$ in which each copy of H_0 contains at most one vertex of $\mathscr{V}$. As a result, the vertices in any copy of H_0 still have reasonably large common degree in A and the König–Hall Theorem still holds. Observe that in the case $|\mathscr{V}| = 0$, the vertices of V_0 were inserted into different H_0's. Since $|\mathscr{V}| < 2\beta/\beta_2 n < \beta_2 n$, we can simply move the vertices in $\mathscr{V}$ to V_0 and then follow the procedure used in the ideal case.

Case 2: $|A| < (\frac{1}{2} - \gamma)n$

In Case 1, each copy of H has one of its w-vertex color class coming from A. When $|A| < (\frac{1}{2} - \gamma)n$, we will change the tiling such that some copies of H have their u-vertex color class chosen from A. The rest is similar to Case 1.

Case 3: $|A| > (\frac{1}{2} - \gamma)n$

Similarly to Special Case (I) in Section 4.4.1, the existence of $|A| - (\frac{1}{2} - \gamma)n$ copies of w-stars in A will help to reduce the $|A|$. We then finish the tiling using an argument similar to Case 1. □

Acknowledgements

The ideas presented in this paper were initiated and refined with much help from Endre Szemerédi, we are grateful for his help. We thank János Komlós for his valuable comments that made this manuscript possible. We also thank two anonymous referees for their numerous suggestions that improved the presentation.

References

[1] S. Abbasi, The solution of the El-Zahar problem, Ph.D. Thesis, Rutgers University, 1998.
[2] N. Alon, R. Yuster, Almost h-factors in dense graphs, Graphs Combin. 8 (1992) 95–107.
[3] N. Alon, R. Yuster, h-Factors in dense graphs, J. Combin. Theory B 66 (1996) 269–282.
[4] K. Corrádi, A. Hajnal, On the maximal number of independent circuits in a graph, Acta Math. Hungar. 14 (1963) 423–439.
[5] G.A. Dirac, Some theorem on abstract graphs, Proc. London Math. Soc. 2 (1952) 69–81.
[6] M.H. El-Zahar, On circuits in graphs, Discrete Math. 50 (1984) 227–230.
[7] A. Hajnal, E. Szemerédi, Proof of a conjecture of Erdös, in: P. Erdös, V.T. Sós (Eds.), Combinatorial Theory and its Applications, II. Colloquia Mathematica Societatis János Bolyai, North-Holland, Amsterdam/London, 1970.
[8] J. Komlós, Tiling Turán theorems, Combinatorica 20 (2000) 203–218.
[9] J. Komlós, G.N. Sárközy, E. Szemerédi, Blow-up Lemma, Combinatorica 17 (1997) 109–123.
[10] J. Komlós, G.N. Sárközy, E. Szemerédi, Proof of the Alon-Yuster conjecture, Combinatorics, Prague, 1998; Discrete Math. 235 (2001) 255–269.
[11] J. Komlós, G.N. Sárközy, E. Szemerédi, An algorithmic version of blow-up lemma, Random Structures Algorithms 12 (1998) 297–312.
[12] J. Komlós, M. Simonovits, Szemerédi's regularity lemma and its applications in graph theory, in: D. Miklós, V.T. Sós, T. Szónyi (Eds.), Bolyai Society Mathematical Studies 2, Combinatorics, Paul Erdős is Eighty, Vol. 2, Keszthely, Hungary, 1993, János Bolyai Math. Soc., Budapest, 1996, pp. 295–352.
[13] E. Szemerédi, Regular partitions of graphs. Colloques Internationaux C.N.R.S No. 260—Problèmes Combinatoires et Théorie des Graphes, 1976, pp. 399–401.